math221.1 0 Applied calculus of two variables

Gunnar Stefansson

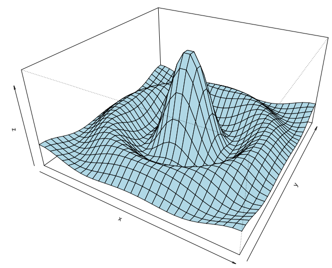
November 12, 2015

Contents

1 On functions of two variables

1.1 Extensions of univariate functions

One can extend the univariate case of $f : \mathbb{R} \to \mathbb{R}$ in several ways.
Here we will consider only the simplest case: $f : \mathbb{R}^2 \to \mathbb{R}$ so $f(x,y) \in \mathbb{R}$.



The function $f(x,y) = sin(x^2+y^2)/(x^2+y^2)$, plotted as a surface of points $(x,y,z) \in \mathbb{R}$ with $z = f(x,y)$.

1.1.1 Details

One can extend the univariate case of $f : \mathbb{R} \to \mathbb{R}$ in several ways.

Here we will consider only the simplest case:

$f : \mathbb{R}^2 \to \mathbb{R}$

so f is of the form $f(x,y) \in \mathbb{R}$.

1.1.2 Examples

Examples:

$f(x,y) = x+y$

$f(x,y) = x^2+y^2$

Another interesting function is $f(x,y) = xy^2$. Note that, as a function of y, the behavior of this function changes completely depending on the sign of x.

$f(x,y) = (x-y)^2$

$f(x,y) = x/y$ if $y \neq 0$

A popular functions is:

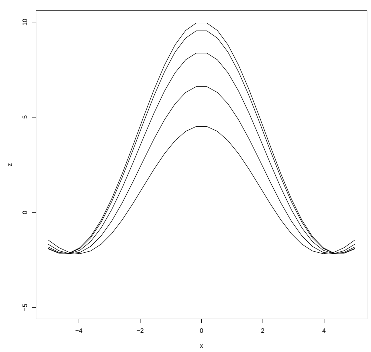
$$f(x,y) = sin(x^2+y^2)/(x^2+y^2)$$

This last function can be drawn with the R commands

```
x <- seq(-10, 10, length= 30)
y <- x
f <- function(x, y) { r <- sqrt(x^2+y^2); 10 * sin(r)/r }
z <- outer(x, y, f)
```

```
z[is.na(z)] <- 1
op <- par(bg = "white")
persp(x, y, z, theta = 30, phi = 30, expand = 0.5, col = "lightblue")
```

1.2 Investigating one variable at a time



Plots of a function $z = f(x,y)$, viewed as a function of x, for several fixed levels of y.

1.3 Contour plots

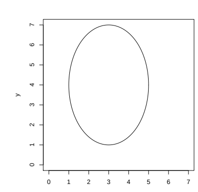
A contour plot is a set of points of the form

$$\left\{(x,y) \in \mathbb{R}^2 : f(x,y) = c\right\}$$

for some number c.

1.4 The equation $F(x,y) = c$

$F : \mathbb{R}^2 \to \mathbb{R}$
Then $F(x,y) = c$ defines a relationship between x and y.
Note that this is a contour of the function.
We can sometimes solve this equation to write y as a function of x.
We can also differentiate this equation...



Points $(x,y) \in \mathbb{R}^2$ satisfying $\frac{(x-3)^2}{4} + \frac{(y-4)^2}{9} = 1$.

1.4.1 Details

The methodology here involves differentiating an equation $F(x,y)=c$, which defines y only implicitly as a function of x.

Any terms in the function are differentiated taking into account that derivatives of components such as $h(y)$ become $\frac{d}{dx}h(y)=h'(y)\frac{dy}{dx}$.

1.4.2 Examples

Example: Consider the function

$$F(x,y)=\frac{(x-3)^2}{4}+\frac{(y-4)^2}{9}$$

and suppose we are interested in the particular contour $F(x,y)=1$.

We can first analyse this contour by noticing that we can write

$$F(x,y)=\left(\frac{x-3}{2}\right)^2+\left(\frac{y-4}{3}\right)^2$$

and it follows that if we write

$$\begin{aligned} u &= \frac{x-3}{2} \\ v &= \frac{y-4}{3} \end{aligned}$$

so we also have

$$\begin{aligned} x &= 2u+3 \\ y &= 3v+4 \end{aligned}$$

then $F(x,y)=1$ is equivalent to $u^2+v^2=1$ so (u,v) must lie on the unit circle and (x,y) are a transformation of (u,v) obtained by stretching and then shifting the circle from $(0,0)$ to $(3,4)$, resulting in an ellipse.

Note that an ellipse does NOT define y as a function of x.

On the other hand, for a given point, (x_0,y_0), on the curve, we can consider y as a function of x in a small neighborhood around the point and write $y=f(x)$ (or $y(x)$ etc). Alternatively one can simply write y but keep in mind that y is now a function of x.

Since $F(x,y)=c$ now defines y as a function of x, we can write $F(x,f(x))=c$ and this is an equation which should hold for x in some interval and this is an equation which we can in principle differentiate.

$$F(x,f(x))=1\Rightarrow\frac{(x-3)^2}{4}+\frac{(f(x)-4)^2}{9}=1\Rightarrow\frac{d}{dx}\left(\frac{(x-3)^2}{4}+\frac{(f(x)-4)^2}{9}\right)=0$$

We need to be careful with the differentiation since one of the terms is a composite function, but we obtain:

$$\frac{2(x-3)}{4}+\frac{2(f(x)-4)}{9}f'(x)=0$$

and this we can rewrite to obtain

$$f'(x)=-\frac{\frac{2(x-3)}{4}}{\frac{2(f(x)-4)}{9}}.$$

We have therefore shown that if (x,y) is a point on the curve where $y \neq 4$, then we can write $y = f(x)$ with

$$f'(x) = -\frac{\frac{2(x-3)}{4}}{\frac{2(y-4)}{9}},$$

in other words we can find the derivative of the function without knowing the shape.

Example:If $xy = \arctan(y)$ then we can write $y = f(x)$ and then find $\frac{dy}{dx}$ by differentiation both sides of the equation and solving for $f'(x)$.

1.5 Partial derivatives

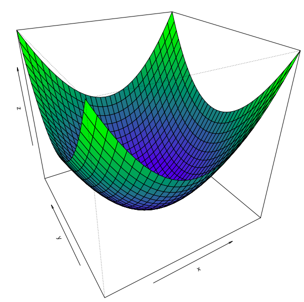
A function of two or more variables can be inspected as a function of one variable at a time:
Suppose $F : \mathbb{R}^2 \to \mathbb{R}$ is differentiable in each variable.
We write $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ for the two derivatives.
A local maximum (or minimum) must have

$$\frac{\partial F}{\partial x} = 0$$

and

$$\frac{\partial F}{\partial y} = 0$$

but this may still not be a maximum or a minimum.



The function $F(x,y) = x^2 + y^2$.

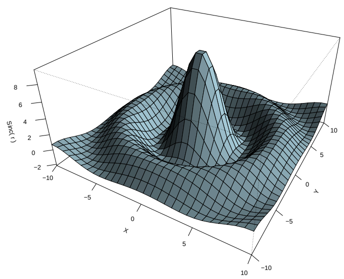
1.5.1 Examples

Example:$F(x,y) = x^2 + y^2$

Example:$F(x,y) = x^2 - y^2$

2 More on real-valued functions of two variables

2.1 Real functions of more than one variable



The function $\sin(x^2+y^2)/(x^2+y^2)$.

2.1.1 Details

The general real-valued function of two (real) variables is a function $f : \mathbb{R}^{\nvDash} \to \mathbb{R}$. This can be defined by any formula which includes the two variables.

The general real-valued function of several (real) variables is a function $h : \mathbb{R}^{\ltimes} \to \mathbb{R}$ defined by some formula.

Definitions of when these functions are continuous are extensions of the univariate case. Loosely, a real-valued function of two real variables is continuous at a point (x_0, y_0) if the values $f(x,y)$ are “close” to $f(x_0, y_0)$ when (x,y) is “close” enough to (x_0, y_0).

The set of points where a function like this is constant:

$$\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c\}$$

is a **level set**, or, in the case of the plane ($n = 2$), a **level curve** or **contour line**.

Values of a function of two variables can be drawn in 3 dimensions, as the set of points

$$\{(xyf(x,y)) : x, y \in \mathbb{R}\}$$

2.1.2 Examples

Example:

$$f : \mathbb{R}^2 \to \mathbb{R} \; f(x,y) = x^2 + y^2$$

Here the contour curves are circles. **Example:**

$$g : \mathbb{R}^3 \to \mathbb{R} \; g(x,y,z) = xyz$$

Example:

$$h : \mathbb{R}^n \to \mathbb{R} \; \ldots$$

2.2 Partial differentiation

2.2.1 Details

In principle, just differentiate with respect to one variable at a time. Write

$$\frac{\partial f(x,y)}{\partial x}$$

$$\frac{\partial f(x,y)}{\partial y}$$

To be differentiable, these partial derivatives need to satisfy criteria...if the partial derivatives are continuous, then the function is differentiable.

2.3 The gradient

If $f : \mathbb{R}^n \to \mathbb{R}$, then we define the **gradient** of f as the vector

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_1} \\ \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_n} \end{bmatrix}$$

2.3.1 Details

If $f : \mathbb{R}^n \to \mathbb{R}$, then we define the **gradient** of f as the vector

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_1} \\ \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_n} \end{bmatrix}$$

2.3.2 Examples

Example: Consider the function $f(x,y) = x^4 + x^2(1-2y) + y^2 - 4x + 4$. The gradient of this function at a general point (x,y) is

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 4x^3 + 2x(1-2y) - 4 \\ 2y - 2x^2 \end{bmatrix}$$

Hence e.g. at $(x,y)=(0,1)$ we can calculate the gradient at this particular point as

$$\nabla f(\mathbf{x}) = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

and we can identify any potential maxima or minima as the points where $\nabla f = \mathbf{0}$, i.e. where both $0 = \frac{\partial f}{\partial x} = 4x^3 + 2x(1-2y) - 4$ and $0 = \frac{\partial f}{\partial x} = 2y - 2x^2$. For this to occur we need $y = x^2$ and also $0 = 4x^3 + 2x(1-2x^2) - 4 = 2x - 4 \Rightarrow x = 2$ and therefore $y = 4$.

2.4 Higher order derivatives

2.4.1 Details

If the functions are differentiable in the coordinates then we can keep on differentiating to get mixed derivatives...

2.4.2 Examples

Example: For a function of only two variables we can compute

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

and

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

Example: Consider the function ...

2.5 The Hessian matrix

2.5.1 Details

The Hessian matrix is the matrix of all combinations of second-order derivatives, for example:

$$H = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial y \partial x} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix}$$

2.5.2 Examples

The Hessian matrix is the matrix of all combinations of second-order derivatives, for example:

$$H = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial y \partial x} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix}$$

Example: Consider the function $f(x,y) = x^4 + x^2(1-2y) + y^2 - 4x + 4$. The gradient of this function at a general point (x,y) is

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x_1} \\ \frac{\partial f(x,y)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x^3 + 2x(1-2y) - 4 \\ 2y - 2x^2 \end{bmatrix}$$

Hence e.g. at $(x,y) = (0,1)$ we can calculate the gradient at this particular point as

$$\nabla f(\mathbf{x}) = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

and the Hessian is

$$H = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial y \partial x} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x^2 + 2(1-2y) & -4x \\ -4x & 2 \end{bmatrix}$$

so e.g. at the point $(x,y) = (0,1)$ the value of the Hessian is ...

3 Maxima and minima of real-valued functions of two variables

3.1 Unconstrained local optimization

Local extrema must satisfy

$$\nabla f(x,y) = 0$$

(if the derivatives exist everywhere)

3.1.1 Details

Local extrema must satisfy

$$\nabla f(x,y) = 0$$

(if the derivatives exist everywhere)

3.1.2 Examples

Example: Consider again the function $f(x,y) = x^4 + x^2(1-2y) + y^2 - 4x + 4$. The gradient of this function at a general point (x,y) is

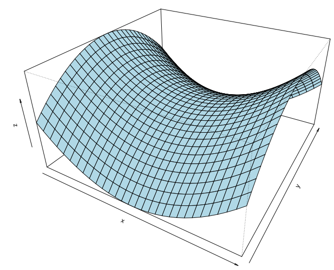
$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x_1} \\ \frac{\partial f(x,y)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x^3 + 2x(1-2y) - 4 \\ 2y - 2x^2 \end{bmatrix}$$

To find potential maxima and minima we solve the equations $\nabla f(\mathbf{x}) = \mathbf{0}$ to find $(x,y) = (2,4)$.

3.2 Classification of extrema

If $\nabla f(x_0,y_0) = 0$, H the Hessian with eigenvalues $\lambda_1 > \lambda_2$.

- $\lambda_1 > \lambda_2 > 0$: local minimum $\Leftarrow det(H) > 0,\ tr(H) > 0$
- $0 > \lambda_1 > \lambda_2$: local maximum $\Leftarrow det(H) > 0,\ tr(H) < 0$
- $\lambda_1 > 0 > \lambda_2$: saddle point $\Leftarrow det(H) < 0$



The function $f(x,y) = x^2 - y^2$.

3.2.1 Details

λ, is an **eigenvalue** a matrix A if there is a non-zero $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$.

Eigenvalues can be found by solving the **characteristic equation**: $\det(A - \lambda I) = 0$

If $\nabla f(x_0,y_0) = 0$, H is the Hessian (of continuous partial derivatives) and

- The two eigenvalues of H are positive, then f has a local minimum at (x_0,y_0); $\Leftarrow det(H) > 0,\ tr(H) > 0$
- The two eigenvalues of H are negative, then f has a local maximum at (x_0,y_0); $\Leftarrow det(H) > 0,\ tr(H) < 0$
- The two eigenvalues of H are of different sign, then f has a saddle point at (x_0,y_0); $\Leftarrow det(H) < 0$

3.2.2 Examples

Example: Consider the function $f(x,y) = x^4 + x^2(1-2y) + y^2 - 4x + 4$. The gradient of this function at a general point (x,y) is

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x_1} \\ \frac{\partial f(x,y)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x^3 + 2x(1-2y) - 4 \\ 2y - 2x^2 \end{bmatrix}$$

Weknow that the only local extremum is (2,4) and and since the Hessian is

$$H = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial y \partial x} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x^2 + 2(1-2y) & -4x \\ -4x & 2 \end{bmatrix}$$

so at the point $(x,y) = (2,4)$ the value of the Hessian is ...

We can now find the eigenvalues at this point by solving the equation $det(H - \lambda I) = 0$ for λ.

3.3 Constrained optimization

To maximize $f(\mathbf{x})$ with respect to $g(\mathbf{x}) = 0$, where both are real-valued,
set up the Lagrange function

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

and solve

$$\frac{\partial L}{\partial x_i} = 0, \ i = 1, \ldots, n$$

along with $g(\mathbf{x}) = 0$.
This will (under certain regularity conditions) give the extrema of f with respect to $g = 0$.

3.3.1 Details

To maximize $f(\mathbf{x})$ with respect to $g(\mathbf{x}) = 0$, where both are real-valued,

set up the Lagrange function

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

and solve

$$\frac{\partial L}{\partial x_i} = 0, \ i = 1, \ldots, n$$

along with $g(\mathbf{x}) = 0$.

This will (under certain regularity conditions) give the extrema of f with respect to $g = 0$.

3.3.2 Examples

Example: Consider the optimization problem to minimize $f(x,y) = x^2 + y^2$ subject to $g(x,y) = x + y - 1 = 0$.

Here the Lagrangian is

$$L(x,y,\lambda) = x^2 + y^2 + \lambda(x + y - 1)$$

and hence

$$0 = \tfrac{\partial L}{\partial x} = 2x + \lambda \Rightarrow \lambda = -2x$$
$$0 = \tfrac{\partial L}{\partial y} = 2y + \lambda \Rightarrow \lambda = -2y$$

from which it follows that the extremum must satisfy $x = y$. Since we also have $x + y = 1$, the only potential local minimumis $x = y = \frac{1}{2}$

3.4 Classification of constrained extrema

Write $L(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$ and suppose $\mathbf{x}^*$ is a potential extremum with $0 = \nabla_{\mathbf{x}^*} L = \nabla f(\mathbf{x}^*) + \lambda^* \nabla g(\mathbf{x}^*)$ and $g(\mathbf{x}^* = 0$.
Further, define the Hessian of L, with respect to $\mathbf{x}$ as

$$H = \nabla^2_{\mathbf{x}^*} L = \nabla^2 f(\mathbf{x}^*) + \lambda^* \nabla^2 g(\mathbf{x}^*)$$

If eigenvalues of H are all positive, then $\mathbf{x}^*$ is a local minimum.

3.4.1 Details

Write $L(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$ and suppose $\mathbf{x}^*$ is a potential extremum with $0 = \nabla_{\mathbf{x}^*} L = \nabla f(\mathbf{x}^*) + \lambda^* \nabla g(\mathbf{x}^*)$ and $g(\mathbf{x}^* = 0$.

Further, define the Hessian of L, with respect to $\mathbf{x}$ as

$$H = \nabla^2_{\mathbf{x}^*} L = \nabla^2 f(\mathbf{x}^*) + \lambda^* \nabla^2 g(\mathbf{x}^*)$$

If eigenvalues of H are all positive, then $\mathbf{x}^*$ is a local minimum.

Note that H is just computed at $\mathbf{x}^*$. It is also true that a much weaker condition is sufficient for the point to be a minimum, but this is outside the scope of these notes.

3.4.2 Examples

Example: For $f(x,y) = x^2 + y^2$ and $g(x,y) = x + y - 1$ we have $L(x,y,\lambda) = x^2 + y^2 + \lambda(x + y - 1)$, $\nabla_{\mathbf{x}} L = (2x + \lambda, 2y + \lambda)'$ and thus

$$\nabla^2_{\mathbf{x}} L = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

which has both eigenvalues equal to two and therefore both positive.